Journal of Applied Horticulture, 23(1): 84-88, 2021
DOI: https://doi.org/10.37855/jah.2021.v23i01.16

Journal of Applied Horticulture
ISSN: 0972-1045

Short-term forecasting of greenhouse tomato price before supply to the market: Isfahan-Iran

M. Ramezani and A. Papzan*

Department of Agricultural Extension and Education, College of Agriculture, Razi University, Kermanshah, Iran.
**E-mail: abdolhamid.papzan@gmail.com*

Abstract

A reliable forecast of the prices of agricultural commodities can help to allocate resources optimally, enhance efficiency and farmer income, and alleviate fluctuations. Given the importance of accurate forecasts, the present paper investigates how one can forecast greenhouse tomato prices at one, two, three, and six-month horizons using different support vector machines and artificial neural network algorithms. The data on variables affecting the price of greenhouse tomatoes were collected through a field study for a short-term period from November 2014 to January 2017. The wholesale price of the crop was drawn from a market study for this period. The results show that the trend forecasted through General Regression Neural Network along with activating function of reciprocal is more efficient to estimate the training data. The Epsilon-SVR SVM acquisition pattern alongside the linear activating function was taken into consideration to estimate the testing data in an efficient way for two-month, three-month or biannual periods.

Key words: Tomato price, price forecasting, artificial neural networks, support vector machines, efficiency comparison

Introduction

Inappropriate distribution of agricultural production, fragmentation of farms, and the lack of coordination and planning pose serious risks to farmers. In fact, excessively short-term fluctuations in crops are negatively associated with production, employment, demand, the use of modern technology, welfare, and social stability. To contribute to planning the production, adjusting the policies, and minimizing the risks affecting authorities and stakeholders, most researchers prefer to model market prices by various statistical methods. Although it has been a complicated, tough challenge to model the forecast of time series prices for a decade, it has been an attractive and interesting topic to study.

In the past, researchers have focused on applying individual models based on one linear or nonlinear forecast technique while some scholars have sought to benefit from statistical methods such as Vector Regression (VAR), Vector Error Correction Model (VECM), Autoregressive Integrated Moving Average (ARIMA), Box-Jenkins algorithms, structural patterns, and Multiple Linear Regression (MLR) to discuss time series data in prices. But, these techniques suffer from drawbacks that have hampered researchers to consider the sophisticated and nonlinear factors affecting the forecast. With the rapid development of artificial intelligence (AI) techniques, artificial neural networks (ANN), fuzzy logic, and machine learning (ML), focus has been drawn to use these techniques to forecast time series data in determining non-linear models. This would allow coping with the constraints of the traditional statistical methods. Much research has been carried out to compare the robustness of time series test results with ANN so far.

Many research works have taken advantage of various ANN models to carry out their studies. Wang *et al.* (2006) used the Hopfield neural networks to predict soya yields. Neural Network Auto-Regressive Model with Exogenous inputs (NNARX) was used by Ahmad *et al.* (2001) to forecast egg price and by Fahimi-Fard *et al.* (2009) to forecast retail price of rice, chicken and egg. However, it seems that the adaptive neuro-fuzzy inference system is benefited to forecast the tomato price (Li *et al.*, 2010), the wholesale price of different types of meat (Feiz-abadi, 2012) and the prices of two crops including rice and barely (Fahimi-Fard *et al.*, 2008; 2009). Overall, the results of several studies such as the forecast of the retail price of chicken (Azarbaijani and Bayari, 2007), the forecast of wheat and cow prices (Asna-Ashari, 2007), the effectiveness of monetary and fiscal policies in agricultural employment (Najafi *et al.*, 2006), and the forecast of price stocks (Ghasemi *et al.*, 2000) illustrate that the accuracy of model prediction of neural network has increased, particularly in ARIMA, when compared to the traditional methods.

Some researchers argue that neural networks are more appropriate to be applied as a strong complement rather than an alternative (Gonzalez, 2000). The investigation carried out by Mortazavi *et al.* (2013) to forecast the weekly retail and wholesale prices of Trout fish, along with the research conducted by Zare-Mehrjerdi *et al.* (2011) to predict meat price show that the traditional model, the ARIMA, is significantly more efficient than the ANN models. A study on forecasting the wholesale price of crops such as tomato, onion, and potato has also demonstrated that neural network seems more convenient in the short run than the other forecast methods while in the long run, there is no evidence of differences in different forecast models (Najafi *et al.*, 2007).

Since, time series of price have both linear and non-linear models in the real world, hybrid or systematically combined models emerged to estimate the linear and nonlinear models simultaneously (Tao *et al.,* 2017).

Certain studies have used hybrid models, including the forecast of monthly prices of chicken (Qarib, 2009), the estimation of the global prices of wheat, corn, and sugar (Moghadasi and Zhale Rajabi, 2013), the price forecast of crops such as soya and rapeseed mustard (Jha and Sinha, 2013), barely price forecast (Zhao, 2017), a short-term prediction of pork price in a combined way based on EEMD, the forecast of vegetable prices (Ye *et al.*, 2016) and the forecast of the future prices for three agricultural crops of wheat, corn and soya. The single models suffer from more errors in price prediction than the combined models.

So far, more than 50 different models of ANN have been designed, implying the widespread and extensive discussion on neural networks. It is, therefore, necessary to compare the efficiency of various traditional models, neural networks, and hybrid methods. The present study evaluates the efficiency of ANN models and SVM regression in 17 ways.

Materials and methods

The present study aimed to explore the feasibility of the short-term forecast of greenhouse tomato price before its supply by greenhouse owners to the market in Isfahan province using ANN and SVM regression. The required data were collected from the field study and were analyzed using different methods of ANN and SVM regression by MATLAB. The data about greenhouse owners' cultivation from 2014 to 2017 was collected by a questionnaire filled by the greenhouse owners and interviewing with them. There are 471 metal greenhouses in Isfahan province. In the present study, the information about the cultivation of all greenhouse owners producing tomatoes over the given three years was collected through the census method. But in practice, a total of 279 individuals were studied.

ANN and SVM regression were employed to analyze the data. The feed-forward network was used to reduce the error and improve forecasting power. The feed-forward network performs far better than the backward network in estimating functions. Since, many greenhouse owners did not have a proper system to record their daily production and sale rates, the sample greenhouses production function was used to determine the production rate at each greenhouse over the studied intervals. For this purpose, the tomato production rate function was first obtained based on the data recorded in the sample greenhouses for a certain cultivation period. To achieve this goal, the data of the greenhouses recording the daily harvests were collected. The tomato production function was estimated using nonlinear regression after removing the noisy data. The tomato production rate of the sample greenhouses at harvest time was estimated to correspond to the following function per 1000 m^2:

$$y = 0.0000004p^3 - 0.001p^2 + 0.0646p + 0.5149 \qquad (1)$$

$R^2 = 0.971$

P= period (every 10 days were considered to form a period)

y= production rate over a short interval

The total production of each greenhouse differs from that of the others for different reasons such as the differences in the environmental and human factors. The total production of each greenhouse was examined by interviews with the greenhouse owner. Although, the greenhouse owners did not record the daily data of their production, they informed the researcher of their total production based on their experience or the analysis of their sale invoices. The total production rate of each individual greenhouse was redistributed using the above function at short 10-day intervals and the production rates at these intervals. For this purpose, at first the number of days and finally the 10-day periods, when each greenhouse has produced tomatoes, were determined, and the production rate at each period was then calculated according to the above function. The length of the tomato production period in the greenhouses of Isfahan province was, on average, six months with a maximum of nine months. The numbers obtained from this calculation were made dimensionless and normalized such that the sum of all numbers would be equal to 1.

$$W_{ij} = \frac{y_{ij}}{\sum y_{ij}} \qquad (2)$$

$$\sum w_i = 1 \qquad (3)$$

The total production rate and area under cultivation at each interval were calculated through the calculation of the production of each greenhouse at 10-day intervals based on the production function of sample greenhouses, and the algebraic addition of production rate to the cultivation area of all greenhouses.

$$\begin{bmatrix} y_1 \\ y_2 \\ . \\ y_n \end{bmatrix} \begin{bmatrix} W_{11} & W_{12} & \dots & W_{1n} \\ W_{21} & W_{22} & \dots & W_{2n} \\ . & . & \dots & . \\ W_{m1} & W_{m2} & \dots & W_{mn} \end{bmatrix} \qquad (4)$$

The daily sale prices of tomato during the study period were collected and the forward-feed network was created through the calculation of the price at each 10-day period.

The database included data on the cultivation by greenhouse growers for a period ranging from November 2014 to July 2017. The time trend was used to analyze the data. According to the applied method, the data ranging from November 2016 to January 2017 were used as the training set while the data for one, two, three and six last months were used as the testing set.

In this study, six learning models of GEN, RBF, MLP, GRNN, CNN, and GMDH were employed to predict the price of neural networks to assess their performance. In the design of the Multilayer Perceptron Neural Network (MLP), a combination of sigmoid and linear stimulus functions were used for hidden and output layer activation functions, which formed four models, to obtain relatively improved results.

In the GRNN model, two Gaussian and reciprocal activation models were used to forecast the tomato price. As a result, ten neural network models were applied to forecast the price. It is important to characterize the parameters and elements of the model, the learning coefficient, the number of hidden layers of the activation functions and the output layer. Various linear and nonlinear functions in the hidden layer and output layer were used in the present study. The applied network is a neural network characterized by error back-propagation, with learning rule of LM, a learning rate of 0.01, and maximum frequencies of 2000. This method serves to provide more accurate results in a shorter time than when a decreasing gradient technique is used.

The Support Vector Regression (SVR) employed two learning patterns of Epsilon-SVR and Nu-SVR with four kernel functions, namely Sigmoid, Polynomial, RBF and linear functions were used. They generated eight different combinations to forecast prices.

The findings are obtained by prediction, while the prediction accuracy is evaluated by measures such as root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) by the following equations where $X_{model,i}$ indicates the amount of forecast, is $X_{obs,\ i}$ the real value, and n is number of observations:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}} \quad (5)$$

$$MAE = \frac{1}{n}\sum_{i}(X_{obs,i} - X_{model,i})^2 \quad (6)$$

Absolute error percentage (APE) can be calculated by the following equation:

$$APE = \frac{\left|(X_{obs,i} - X_{model,i})^2\right|}{X_{obs,i}} * 100 \quad (7)$$

The mean absolute percentage error (MAPE) is computed as follows:

$$MAPE = \left(\frac{1}{N}\right)\sum_{N} APE \quad (8)$$

Where, N indicates the number of prediction period. If the above statistics are close to zero, the best measure will be provided.

Results

In this section, firstly, the individual characteristics of the respondents are considered (Table 1). Based on descriptive findings, gender of 91.8 % of greenhouse growers is male. It seems that greenhouse grower job is considered to be a men's job. During the interviews, it is also found that the percentage of women which themselves have been trying to establish a greenhouse is less than that. Women have mostly established greenhouses with the help of their spouse, father and.or their brother.

The mean age of the individuals of the study is 44.29 years and the standard deviation of the respondents' age is 12.38. The youngest respondent is 23 years old and the oldest respondent is 81 years old. About 51.1 % of the members are under 40 years old and 31.9 % of the members are over 50 years old. About 16.8 % of the population is comprised of individuals aged 41-50 years. The majority of the age group of the greenhouse growers is made up of young people.

41.9 % of the members of the statistical population have university education, about 27.6 % have middle school education and lower education. About 30.5 % of the sample population have high school education and diploma. On average, the level of education of the members is limited to the associate degree.

The research findings indicate that 45.5 % of greenhouses in the province have an area of smaller than 4000 m^2 and are economically small, even though 37.3% have a moderate scale. Only, 17.2% have an area of above 8 acres and are considered to be large and commercial greenhouses. The smallest greenhouse is 500 m^2 while the largest one is 8 hectares.

Respondents were asked to list one to four substantial problems. The findings of the study show that 60.7 % of the greenhouse growers' problems are related to economic factors. Specifically,

Table 1. Characteristics of the respondents

Variable /Level of evaluation	Frequency	Percent	Cumulative percent
Gender			
Female	23	8.2	8.2
Male	256	91.8	100
Total	279	100.0	
Age			
21-30	35	12.5	12.5
31-40	108	38.7	51.3
41-50	47	16.8	68.1
51-60	63	22.6	90.7
61-70	24	8.6	99.3
71 and above	2	0.7	100
Total	279	100	
Education			
Illiterate	3	1.1	1.1
Elementary & Guidance School	74	26.5	27.6
High School & Diploma	85	30.5	58.1
Technical Diploma	16	5.7	63.8
BS	72	25.8	89.6
MS & Upper Degrees	29	10.4	100
Total	279	100	
Greenhouse area (1000 m^2)			
Less than 1	1	0.4	0.4
from 1 to 2	4	1.4	1.8
from 2 to 3	37	13.3	15.1
from 3 to 4	98	35.1	50.2
from 4 to 5	64	22.9	73.1
from 5 to 6	25	9	82.1
from 6 to 7	13	4.7	86.7
from 7 to 8	10	3.6	90.3
from 8 to 10	13	4.7	95
More than 10	14	5	100
Total	279	100	

the economic factor and the price fluctuations are the main concern of the greenhouse growers.

Price forecasting: The relationship between the production amount and its market price was first evaluated by a chart in MS-Excel software before they were forecasted on the basis of the neural network. Fig. 1 illustrates the results of this evaluation. According to the results, a reverse relationship was observed between the market prices of crops and their supply rate. Interestingly, the greenhouse growers responded to market prices during a two-month time period and they reduced the cultivation area and the production as well. This result shows that responding to market prices in greenhouses is much faster than the traditional farming systems and prices influence cultivation effectively. On the other hand, an approximately four-month period can be observed for reaching from the minimum tomato price to its maximum. The tomato prices showed a fluctuation rate of about 700 %, which is excessively high. The maximum price occurred from October to January when the traditional cultivation was limited, giving rise to reasonably higher prices.

Diverse models of neural networks were applied to evaluate the accuracy of forecasts and the efficiency of calculations.

The value of statistics, RMSE, MSE, MAE, and MAPE were

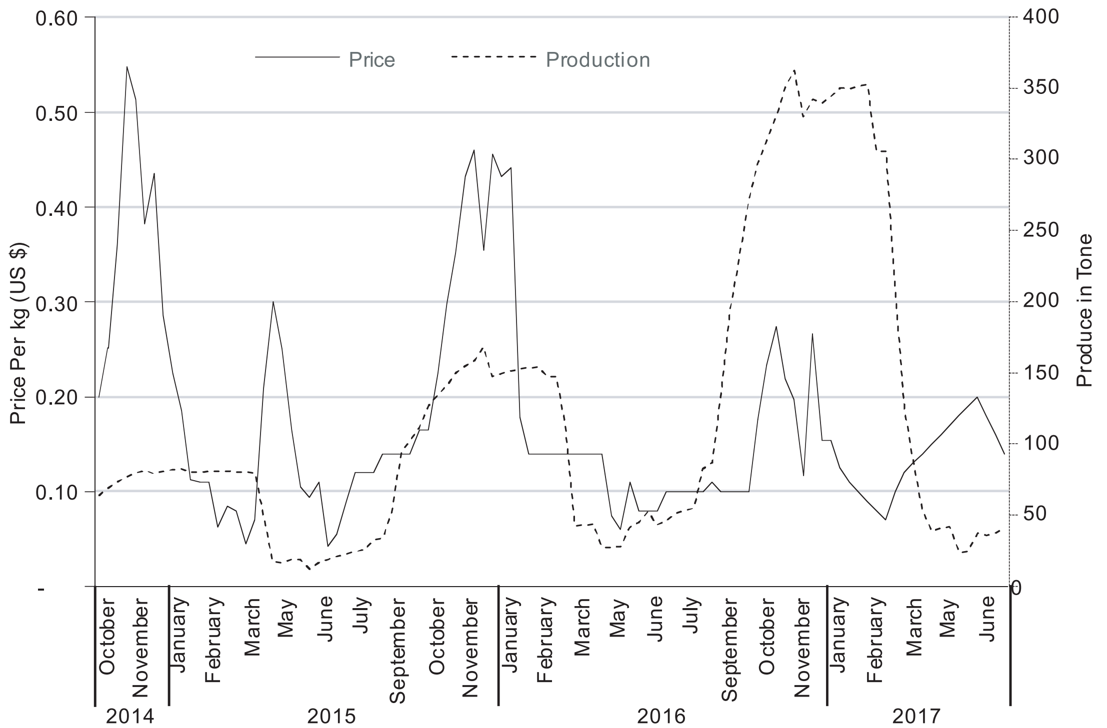

Fig.1. The evaluation of the association between tomato production and its market prices during 2014-2017.
Source: Result of Analysis of Field Survey Data, 2017

assessed by the training and testing tests for each of the above models over one, two, three and six months. The results are summarized in Table 2.

The applied GRNN (reciprocal) was more efficient, particularly in the training set, than any other neural network models to make assessments for a one, two, three and six-month period while the SVM $_{\text{Epsilon-SVR(Sigmoid)}}$ and SVM $_{\text{Nu-SVR(Linear)}}$ were more efficient in the testing set for a one and six, two and three- month period, respectively.

Discussion

The results revealed that when the planting type and time of the individual greenhouses are known, it will be possible to forecast crop prices after their supply to the market. This will help in the management of the market. The importance of short-term forecasts and management of the market and its role in farmers' life cannot be exaggerated. Although, forecasts in Iran are influenced by diverse factors including political instability, inappropriate economic structure, uncertain economic relationships, and the decline of the value of national exchange, the sound pre-planting forecast of crop prices can go a long way to improve farmers' livelihood in the short run even under these unreliable conditions.

The present study was conducted to compare the performance of different neural network models to predict the prices in the short run. The findings suggest that various models of neural networks are different from others in terms of their performance. In the light of different findings from neural networks to predict, it is recommended to researchers, planners, and policymakers to take advantage of various neural network models to predict the agricultural products prices.

Regarding the analyzed data, the GRNN (reciprocal) model can predict appropriately in the training sets of different periods, and SVM is capable of predicting the testing sets from one-month to

Table 2. The assessment of the accuracy of different models of neural networks in forecasting tomato prices in a one, two, three and six-month period

Period	Model	Accuracy measure for training set			Accuracy measure for testing set		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE
Monthly	GRNN$_{(reciprocal)}$	6.67	2.94	0.31	21.19	18.07	1.21
	SVM $_{\text{Epsilon-SVR(Sigmoid)}}$	112.56	84.28	5.93	16.86	14.51	0.94
Two-month	GRNN$_{(reciprocal)}$	6.68	2.88	0.32	51.35	44.49	2.43
	SVM $_{\text{Nu-SVR(Linear)}}$	114.3	86.76	6.13	19.8	16.44	1
Three-month	GRNN$_{(reciprocal)}$	6.79	2.97	0.33	46.59	38.94	2.17
	SVM$_{\text{Nu-SVR(Linear)}}$	116.01	88.25	6.18	18.43	15.88	0.95
Six-month	GRNN$_{(reciprocal)}$	7.12	3.18	0.35	67.55	55.74	4.88
	SVM $_{\text{Epsilon-SVR(Sigmoid)}}$	127.6	97.92	6.17	46.97	36.8	2.7

Source: Result of Analysis of Field Survey Data, 2017. All RMSE and MAE values were multiplied by 10^{-3}.

six-month periods. Thus, they are accurate tools to predict the economic variables along with other methods.

Adaptive comparison of a monthly to six-month periods suggests that the prediction error has increased by improving their duration in the testing sets. The estimated prices on the basis of the testing and training sets, also, suggest that the majority of errors related to the price forecast have occurred for the maximum or minimum one.

If the significance of the study is so that only one forecast model is intended to be used and there is limited data, then it would be better to make use of the GRNN (reciprocal) models. The results show that greenhouse tomato prices fluctuate greatly in the short run, so some policies should be adopted by the government based on the results of the forecast of the short-term fluctuations of prices in order to balance the market before the crops are supplied to the market and effective actions should be adopted to support the farmers as well. It is worthwhile to mention that the results of price prediction can help farmers to determine the appropriate time for the supply of the crops to the market, as well.

References

Ahmad, H. A., G.V. Dozier and D.A. Roland, 2001. Egg price forecasting using neural networks. *The Journal Applied Poultry,* 11: 162-171.

Asna-Ashari, H. 2007. *The Impact of Monetary and Fiscal Policies on Employment of Agricultural Sector*, Unpublished MSc. thesis, Department of Agricultural Economic, Zabol University, 2007.

Azarbaijani, K. and L. Bayari, 2007. Energy Consumption Forecasting Based on Sustainable Development: Case Study of Iran. *The 1st Conference on Challenge and Prospects of Sustainable Developmen,* Isfahan-Iran, 2007, p. 38-50.

Fahimi-Fard, S.M., M. Salarpour and M. Sabohi, 2009. Applying the ANFIS algorithm compared with econometrics ARIMA model to predict the retail prices of agricultural products. *Agricultural Economics Journal*, 2: 183-165.

Fahimi-Fard. S.M., A.L., Keikha and M. Salarpour, 2008. Predicting the prices of selected crops by using the integrated neural network-autoregressive method with exogenous inputs (NNARX). *The Economic Agricultural Development Research,* 2: 54-46.

Feiz-abadi, Y. 2012. Comparison the ANFIS algorithm compared with econometrics ARIMA model to predict the wholesale price of veal, *The 1st Conference on Agriculture and Sustainable Resources,* Tehran, Iran, 2012, p. 100-120.

Ghasemi, A., H. Asadpour and M. Shahsadeghi, 2000. The application of neural network in time series forecasts and its comparison with ARLMA model. Iranian Journal Trade Studies, 14: 87-120.

Gonzalez, S. 2000. Neural Networks for Macroeconomic Forecasting: A Complementary Approach to Linear Regression Models. Department of Finance, University of Ottawa.

Jha, G.K. and K. Sinha, 2013. Agricultural price forecasting using neural network model: an innovative information delivery system. Agricultural Economics Research Review, 26: 229-239.

Li G.Q., S.W. Xu and Z.M. Li, 2010. Short-term price forecasting for agro-products using artificial neural networks. Agriculture Agricultural Science Procedia, 1: 278-287.

Moghadasi, R. and M. Zhale Rajabi, 2013. Forecasting power comparison of hybrid and conventional models (case study of wheat, corn and sugar world prices), Journal Agricultural Economics Research, 5: 1-22.

Mortazavi, S., S. Hasanlou and B. Alaie, 2013. Predicting the retail and wholesale prices of crops. Economy and Agricultural Development, 25: 364-355.

Najafi, B., M. Zibaii, M.H. Sheikhi and M.H. Tarzkar, 2007. Predicting the Prices of certain farming products in fars province: artificial neural network apply. The Water Soil Science Journal, 11: 501-512.

Najafi, B. and M.H. Tarzkar, 2006. Forecasting of Iran pistachio export: application of artificial neural network. Journal Trading Bulletin. 29: 191-214.

Qarib, F. 2009. Forcasting Price of Chicken by Using Neural Network and Genetic Algorithm. Unpublished M.Sc. thesis, Department of Agricultural Economic, Tarbiat Modarres university, 2009.

Tao, X., L.I. Chongguang and B. Yukun, 2017. An improved EEMD-based hybrid approach for the short-term forecasting of hog price in China. Agricultural Economics, 63: 136-148.

Wang, l.G., Qi. Fu and Y. Liu, 2006. Soybean yield forecast application based on Hopfield ANN model. The Journal American Science, 3: 85-89.

Yu, L., Z. Wang and L.A. Tang, 2015. A decomposition–ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting. Applied Energy, 15: 251-267.

Ye, L., X. Qin Li, Y. Liu and W. Liang, 2016. Vegetables price forecasting in Hainan province based on linear and nonlinear combination model. The 13th International Conference on Service Systems and Service Management, Kunming, China, 2016, p. 1-5.

ZAre-Mehrjerdi, M. and S. Negarchi, 2011. Comparing the accumulated autoregressive moving average, the fuzzy regression and the accumulated fuzzy regression to predict the price field study: meat price), Agricultural Economic Journal, 21: 100-108.

Zhao, Y, X. Zhang, Z. Shi and L. He, 2017. Grain price forecasting using a hybrid stochastic method. Asia-Pacific Journal Operational Research, 34: 175-200.

Received: January, 2020; Revised: May, 2020; Accepted: June, 2020